



MINISTERO DELLO SVILUPPO ECONOMICO
DIREZIONE GENERALE PER LA LOTTA ALLA CONTRAFFAZIONE
UFFICIO ITALIANO BREVETTI E MARCHI

DOMANDA NUMERO	102001900913769
Data Deposito	06/03/2001
Data Pubblicazione	06/09/2002

Priorità	09/524.401
Nazione Priorità	US
Data Deposito Priorità	

Sezione	Classe	Sottoclasse	Gruppo	Sottogruppo
G	06	F		

Titolo

SISTEMA E PROCEDIMENTO PER GESTIRE RISORSE DI MEMORIZZAZIONE IN UN AMBIENTE DI ELABORAZIONE A GRAPPOLO.
--

DESCRIZIONE dell'invenzione industriale dal titolo:

"Sistema e procedimento per gestire risorse di memorizzazione in un ambiente di elaborazione a grappolo"

di: Dell Products L.P., nazionalità statunitense,
One Dell Way, Round Rock, Texas 78682-2244 - U.S.A.

Inventori designati: Ahmad H. Tawil; Nam V. Nguyen

Depositata il: 6 marzo 2001 TO 2001A 000196

CAMPO TECNICO DELLA DESCRIZIONE

La presente descrizione riguarda in generale il campo dei sistemi di memorizzazione dati e, più particolarmente, un sistema ed un procedimento per gestire risorse di memorizzazione in un ambiente di elaborazione a grappolo.

SFONDO DELLA DESCRIZIONE

Le reti di area di memoria (SAN) includono spesso una raccolta di risorse di memorizzazione di dati accoppiate in modo da comunicare con una pluralità di nodi come stazioni di lavoro e server. Nella presente descrizione, il termine "nodo" e "server" sono utilizzati in modo interscambiabile, con l'intesa che "server" è un tipo di "nodo".

In una SAN, un server può accedere ad una risorsa di memorizzazione dati attraverso una struttura che utilizza il protocollo Fibre Channel.

BUZZI, NOTARO &
ANTONIELLI D'OUX
s.r.l.

Il protocollo Fibre Channel può agire come uno strato fisico comune che permette il trasporto di strati di protocollo superiori multipli, come il protocollo d'interconnessione di sistemi di piccoli computer (SCSI). In un ambiente SAN, il protocollo SCSI può assegnare numeri di unità logica (LUN) alla raccolta di risorse di memorizzazione dati. Il LUN può permettere ad un server in una SAN di accedere a risorse di memorizzazione dati specifiche facendo riferimento ad un LUN SCSI per una risorsa di memorizzazione dati specifica.

Sebbene un sistema di memorizzazione Fibre Channel possa offrire una grande quantità di capacità di memorizzazione, il sistema può anche essere molto costoso da implementare. Di conseguenza, gli utilizzatori spesso cercano di utilizzare in comune la memoria disponibile fornita dal sistema fra server multipli. Purtroppo, se un server accoppiato ad una data SAN utilizza il sistema operativo MICROSOFT WINDOWS NT, il server può cercare di appropriarsi di qualunque LUN visibile per il server. Ad esempio, se un particolare server rileva diversi LUN quando il server viene avviato, si può assumere che ciascun LUN sia disponibile per l'impiego. Quindi, se server WINDOWS NT multipli sono collegati ad un insieme di

BUZZI, NOTARO &
ANTONIELLI D'OUX
s.r.l.

memorie o a risorse di memorizzazione di raccolta dati, ciascun server può tentare di ottenere il controllo di ciascun LUN nell'insieme di memorie. Questa situazione può portare a conflitti quando più di un server prova ad accedere allo stesso LUN.

Un utilizzatore che cerca di risolvere questo problema può ripartire o suddividere in zone la memoria disponibile attraverso il filtraggio o attraverso l'impiego di programmi di gestione mini-porta che presentano caratteristiche di mascheratura LUN. In effetti, questa suddivisione può impedire ad un server che utilizza WINDOWS NT di vedere caratteristiche di memorizzazione che non gli siano assegnate. Questo approccio può essere efficace per server isolati, ma l'approccio presenta diversi inconvenienti in un ambiente di elaborazione a grappolo.

La connessione a grappolo coinvolge la configurazione di un gruppo di server indipendenti in modo che essi compaiono sulla rete come un'unica macchina. Spesso, i grappoli vengono gestiti come una sistema unico, si suddividono uno spazio nominale comune, e sono progettati specificamente per tollerare errori di componenti e per sopportare l'addizione o la sottrazione di componenti in modo trasparente. Purtroppo, siccome un grappolo può

BUZZI, NOTARO &
ANTONIELLI D'OUX
s.r.l.

presentare due o più server che appaiono essere una macchina unica, le tecniche di suddivisione sopra menzionate possono risultare una soluzione inefficace per evitare conflitti quando i due o più server tentano di accedere allo stesso LUN.

Il SERVER A GRAPPOLO MICROSOFT (MICROSOFT CLUSTER SERVER - MSCS) rappresenta una delle tecniche correntemente disponibili per l'arbitraggio di conflitti e la gestione della proprietà di dispositivi di memorizzazione in un ambiente di elaborazione a grappolo. Un sistema MSCS può operare in un grappolo che presenta due server, il server A, che può comandare, ed il server B. Nel funzionamento, il server A può passare un segnale periodico di pulsazione al server B per fare in modo che il server B sappia che il server A è "vivo". Se il server B non riceve una pulsazione tempestiva dal server A, il server B può cercare di determinare se il server A è funzionante e/o se il server B può assumere la proprietà di qualunque LUN riservato al server A. Purtroppo, questi sistemi MSCS possono utilizzare ripristini di bersaglio SCSI durante questo processo, ed i ripristini SCSI possono creare diversi problemi. Ad esempio, un tipico ripristino SCSI nel sistema MSCS può provocare l'interruzione dei processi input/output "I/O"

BUZZI, NOTARO &
ANTONIELLI D'OUX
s.r.l.

pendenti per tutti i server in un dato sistema Fibre Channel. Questi processi I/O interrotti possono eventualmente essere completati ma non fino a quando il bus non è stabilizzato. Questo approccio di interruzione/attesa/ripetizione può avere un effetto dannoso sulle intere prestazioni di sistema.

Oltre a questo effetto potenziale sulle prestazioni, il sistema MSCS ed il suo impiego di ripristino SCSI possono avere un effetto dannoso sull'intera affidabilità di sistema. Nel funzionamento, il sistema MSCS può rispondere soltanto ad un ripristino SCSI per volta. L'impossibilità di rispondere a successivi ripristini SCSI può portare ad un comportamento inatteso e diminuire l'affidabilità di sistema.

SOMMARIO DELLA DESCRIZIONE

Secondo la presente descrizione, sono descritti un sistema ed un procedimento per gestire risorse di memorizzazione in un ambiente di elaborazione a grappolo che forniscono vantaggi significativi rispetto alle tecniche sviluppate precedentemente. Il sistema ed il procedimento descritti possono permettere una gestione di risorse di memorizzazione e di arbitraggio di conflitti con un impatto ridotto sui ripristini SCSI.

Secondo un aspetto della presente

BUZZI, NOTARO &
ANTONIELLI D'OUIX
s.r.l.

descrizione, un procedimento per la gestione di risorse di memorizzazione in un ambiente di elaborazione a grappolo può includere il mantenimento di una riserva su di una risorsa di memorizzazione per un primo nodo dell'ambiente di elaborazione a grappolo. Il nodo può essere, ad esempio, un server, una stazione di lavoro, o qualunque altro dispositivo di elaborazione incluso nel grappolo.

Può essere eseguito uno scollegamento di processo di un terzo per il primo nodo e la riserva mantenuta per il primo nodo può essere rilasciata. In una forma d'attuazione, lo scollegamento di processo di un terzo può verificarsi in risposta ad un comando di scollegamento inviato per conto del primo nodo. Il comando di scollegamento di processo di un terzo può essere inviato, ad esempio, da un secondo nodo o da un commutatore Fibre Channel. Il comando di scollegamento di processo di un terzo può includere informazioni di identificazione che identificano il primo nodo come mittente del comando di scollegamento anche se il primo nodo non era il vero mittente. Le informazioni di identificazione possono includere, ad esempio, un nome world wide ed un identificatore di sorgente assegnato al primo nodo.

BUZZI, NOTARO &
ANTONIELLI D'OUX
s.r.l.

La gestione di risorse di memorizzazione in un ambiente di elaborazione a grappolo può comportare inoltre la suddivisione in zone di un sistema di memorizzazione Fibre Channel. Una zona può raggruppare un primo nodo con un secondo nodo ed una pluralità di risorse di memorizzazione come unità di disco fisso ed altri dispositivi di memorizzazione dati. Nel sistema suddiviso a zone, un secondo nodo può esso stesso scollegarsi dopo che un comando di scollegamento di processo di un terzo è stato emesso per un primo nodo. Dopo che i due nodi sono scollegati, può essere iniziato un ripristino di collegamento con un protocollo di inizializzazione a ciclo (LIP), può essere generata una notificazione di cambiamento di Stato, e può essere ricollegato qualunque nodo di funzionamento.

Secondo un altro aspetto della presente descrizione, un procedimento per la gestione di risorse di memorizzazione in un ambiente di elaborazione a grappolo può includere il ricevimento di un comando di riserva che cerca di riservare una risorsa di memorizzazione per un nodo nel grappolo. In risposta al comando di riserva, può essere emesso un comando d'uscita di riserva persistente SCSI con un'azione di servizio di riserva per riservare la risorsa di memorizzazione per il nodo. Questa

BUZZI, NOTARO &
ANTONIELLO & C.
S.R.L.

riserva persistente può trattenere una riserva cancellabile sulla risorsa di memorizzazione. In una forma d'attuazione, la riserva può essere cancellata emettendo un comando d'uscita di riserva persistente SCSI con un'azione di servizio di cancellazione. I comandi di riserva persistente possono permettere che riserve LUN vengano rilasciate individualmente come opposte alla cancellazione di diverse riserve LUN contemporaneamente ad un ripristino SCSI.

Secondo un altro aspetto della presente descrizione, un sistema di computer azionabile per gestire risorse di memorizzazione in un ambiente di elaborazione a grappolo può includere un primo nodo, un secondo nodo, ed un motore di gestione di risorse azionabile per scollegare il primo nodo in risposta all'insuccesso da parte del secondo nodo di ricevere un segnale di pulsazione tempestivamente dal primo nodo. Il sistema può anche includere mezzi leggibili tramite computer che memorizzano il motore di gestione di risorse ed un'unità di elaborazione centrale accoppiata in modo comunicativo al mezzo leggibile da computer ed azionabile per eseguire il motore di gestione di risorse.

In una forma d'attuazione, il sistema può anche includere una pluralità di piattaforme di elaborazione accoppiate in modo comunicativo al

BUZZI, NOTARO &
ANTONIELLI D'OUX
s.r.l.

primo nodo. Queste piattaforme di elaborazione possono essere, ad esempio, un insieme di personal computer collegati in rete. Il sistema può anche includere un commutatore Fibre Channel accoppiato in modo comunicativo con un primo nodo e con una pluralità di risorse di memorizzazione. Il commutatore Fibre Channel può, in alcune forme d'attuazione, includere un'unità di elaborazione centrale azionabile per eseguire un motore di gestione di risorse.

Un sistema ed un procedimento che incorporano gli insegnamenti della presente descrizione possono fornire miglioramenti significativi rispetto alle soluzioni di gestione di risorse a grappolo convenzionali. Ad esempio, le tecniche descritte possono essere azionabili per meglio gestire ed arbitrare conflitti di risorse di memorizzazione. Come discusso sopra, un ripristino SCSI in un ambiente di elaborazione a grappolo può provocare l'inizio di un approccio di interruzione/attesa/ripetizione per diversi processi I/O, che possono avere un effetto dannoso sulle intere prestazioni di sistema. Gli insegnamenti della presente descrizione possono aiutare a ridurre l'impatto sui ripristini SCSI e sui decadimenti di prestazioni risultanti.

BUZZI, NOTARO &
ANTONIELLO D'OUX
S.r.l.

Inoltre, gli insegnamenti della presente descrizione possono favorire l'evitare problemi di affidabilità di sistema associati a ripristini SCSI in un ambiente di elaborazione a grappolo. Un sistema di risorse a grappolo convenzionale, come un MSCS, può essere incapace di tener conto di ripristini SCSI iniziati durante il disturbo di bus di un ripristino SCSI precedente. Questa limitazione può portare ad un comportamento inatteso e diminuire l'affidabilità di sistema. Poiché gli insegnamenti della presente descrizione possono favorire l'evitare almeno alcuni ripristini SCSI, può essere migliorata l'affidabilità di sistema.

Altri vantaggi tecnici risulteranno evidenti ad un normale esperto del settore in vista della descrizione, delle rivendicazioni, e dei disegni.

BREVE DESCRIZIONE DEI DISEGNI

Una comprensione più completa della presente descrizione e dei suoi vantaggi potrà essere acquisita con riferimento alla descrizione che segue considerata insieme ai disegni annessi, in cui numeri di riferimento simili indicano caratteristiche simili, ed in cui:

la figura 1 rappresenta un diagramma di componenti di una rete di area di memoria che include una forma d'attuazione di un motore di

BUZZI, NOTARO &
ANTONELLI D'OUX
S.p.A.

gestione di risorse che incorpora gli insegnamenti della presente descrizione;

la figura 2 illustra un diagramma di flusso per una forma d'attuazione di un procedimento di gestione di risorse di memorizzazione in un ambiente di elaborazione a grappolo secondo gli insegnamenti della presente descrizione; e

la figura 3 illustra un diagramma di flusso per un'altra forma d'attuazione di un procedimento di gestione di risorse di memorizzazione in un ambiente di elaborazione a grappolo secondo gli insegnamenti della presente descrizione.

DESCRIZIONE DETTAGLIATA DELL'INVENZIONE

La figura 1 rappresenta un diagramma a blocchi generale di una rete di area di memoria (SAN), indicata generalmente con 10. La SAN 10 include due sistemi di elaborazione a grappolo, i grappoli 12 e 14. Come rappresentato, il grappolo 12 include un nodo 16 ed un nodo 18, ed il grappolo 14 include un nodo 20 e 22. I nodi 16, 18, 20, e 22 possono essere, ad esempio, server, stazioni di lavoro, o altri dispositivi di elaborazione di rete. Come rappresentato nella figura 1, il grappolo 12 può sopportare una pluralità di dispositivi clienti come personal computer clienti rappresentati ad esempio con 24.

BUZZI, NOTARO &
ANTONIELLI D'OUX
s.r.l.

La SAN 10 può anche includere un insieme di memoria 26, che può includere, ad esempio, una pluralità di dispositivi fisici di memoria come un'unità di disco rigido sotto il controllo di controllori di memoria ed accoppiati a uno o più di essi. I dispositivi fisici di memoria dell'insieme di memoria 26 possono essere assegnati ad un LUN. Alcuni dispositivi fisici di memoria possono essere raggruppati in volumi RAID essendo assegnato a ciascun volume un unico indirizzo LUN SCSI. Altri dispositivi fisici di memoria possono essere assegnati individualmente ad uno o più LUN. Tuttavia quando i LUN sono assegnati, i LUN della figura 1 possono mappare la memoria fisica disponibile dell'insieme di memoria 26 in una pluralità di dispositivi logici di memoria e permettere a questi dispositivi logici di memoria di essere identificati ed indirizzati.

Nel funzionamento, i nodi 16, 18, 20, e 22 possono comunicare con l'insieme di memoria 26 e trasferire dati a e da esso attraverso la struttura 28 utilizzando il protocollo fibre channel. Come rappresentato nella figura 1, i nodi 16 e 18 possono essere raggruppati nella zona 30 con LUN_1 e LUN_2. Similmente, i nodi 20 e 22 possono essere raggruppati nella zona 32 con LUN_3, LUN_4, e LUN_5.

BUZZI, NOTARO &
ANTONIELLI D'OUX
s.r.l.

L'impiego della suddivisione a zone mediante commutatore per creare la zona 30 può impedire che i nodi 16 e 18 vedano i nodi 20 e 22. Oltre alla suddivisione a zone, la forma d'attuazione della figura 1 può utilizzare la mascheratura LUN. La mascheratura LUN può schermare un nodo specifico o un grappolo rispetto alla visione di determinati LUN. Ad esempio, la mascheratura LUN può impedire ai nodi 16 e 18 di vedere LUN_3, LUN_4, e LUN_5.

Nella forma d'attuazione della figura 1, i nodi 16, 18, 20, e 22 possono essere assegnati ad un unico nome world wide (WWN), che può essere un identificatore a otto byte. L'Istituto di Ingegneria Elettronica (IEEE) assegna blocchi di WWN a produttori in modo tale che i produttori possano costruire dispositivi fibre channel con WWN unici. A scopo illustrativo, nella forma d'attuazione della figura 1, il nodo 16 può avere un WWN "AAA", il nodo 18 può avere un WWN "BBB", il nodo 20 può avere un WWN "CCC", ed il nodo 22 può avere un WWN "DDD". Come tali, i nodi 16, 18, 20, e 22 possono essere identificabili unicamente da altri dispositivi accoppiati alla struttura 28.

I nodi 16, 18, 20, e 22 possono avere informazioni di identificazione oltre ai loro rispettivi WWN. Ad esempio, secondo il protocollo

BUZZI, NOTARO &
ANTONIELLI D'OLIX
s.r.l.

fibre channel, quando un nodo come il nodo 16 è inizializzato e si collega alla struttura 28, al nodo viene assegnato un ID di fibre channel. Questo ID può essere soggetto a cambiamenti ogni volta che si verifica qualche evento di inizializzazione, ad esempio, quando un altro nodo o dispositivo si collega alla struttura 28.

Come rappresentato nella figura 1, la struttura 28 ha ID fibre channel assegnati come segue: il nodo 16 è S_ID_1, il nodo 18 è S_ID_2, il nodo 20 è S_ID_3, e il nodo 22 è S_ID_4.

Nella forma d'attuazione della figura 1, i vari WWN e ID di fibre channel possono essere memorizzati in un mezzo leggibile tramite computer 34, che può essere accessibile per i dispositivi della SAN 10. Come illustrato nella figura 1, la SAN 10 può includere un dispositivo di elaborazione 38 per stabilire la struttura 28. Un tale dispositivo di elaborazione può includere una CPU accoppiata in modo comunicativo al mezzo leggibile tramite computer 34. Un commutatore 36 può anche presentare almeno una porta 40 per l'interfacciamento con altri dispositivi per formare una rete globale fibre channel.

In una forma d'attuazione di un sistema che incorpora gli insegnamenti della presente

BUZZI, NOTARO &
ANTONIELLI D'OUX
s.r.l.

descrizione, il dispositivo di elaborazione 38 può essere azionabile per eseguire un motore di gestione di risorse, che può essere memorizzato in un mezzo leggibile tramite computer 34. Il motore di gestione di risorse può essere azionabile per eseguire diverse funzioni. Ad esempio, il motore di gestione di risorse può essere azionabile per accedere ad una lista mantenuta dei WWN e degli ID fibre channel dei dispositivi SAN 10. Inoltre, il motore di gestione di risorse può essere azionato per riconoscere un comando di ripristino SCSI emesso da un nodo e per convertire il comando in un comando di rilascio di risorse di memoria. Il comando di rilascio di risorse di memoria può essere, ad esempio, una sconnessione di processo di un terzo o un comando di fuori riserva persistente SCSI con un'azione di cancellazione.

In un tipico grappolo MSCS, un comando di ripristino SCSI può essere emesso quando un nodo come i nodi 18 o 20 non riesce a riconoscere il ricevimento di una pulsazione tempestiva 42 o 44 da un rispettivo grappolo compagno. Le pulsazioni 42 e 44 possono permettere che i nodi 18 e 22 rispettivamente "vedano" se i loro grappoli compagni sono ancora funzionanti.

Se, ad esempio, il nodo 18 non può più

"vedere" il nodo 16, il nodo 18 può cercare che venga rilasciata qualunque riserva LUN mantenuta per il nodo 16. Per realizzare questo rilascio, il nodo 18 può inviare un comando di ripristino SCSI per iniziare un ripristino di bus a livello basso dei bus SCSI associati ai nodi 16 e 18. In alcuni sistemi, ad esempio una sistema MSCS, il nodo 18 può attendere per una quantità specificata di tempo prima di provare a riservare il LUN che è stato riservato dal nodo 16. L'attesa permette che il nodo 16 guadagni nuovamente il controllo dei LUN riservati ad esso prima del ripristino SCSI. Come tale, se il nodo 16 è "vivo" nonostante l'insuccesso del nodo 18 nel ricevere la pulsazione 42, il nodo 16 può essere in grado di ristabilire le sue riserve di risorse e nel fare ciò permette che il nodo 18 sappia che esso è "vivo".

Purtroppo, come detto sopra, un ripristino SCSI in un ambiente di elaborazione a grappolo può avere un effetto dannoso sulle prestazioni di sistema e sull'affidabilità di sistema. Il motore di gestione di risorse e di sistema descritto può favorire il limite di affidabilità di un ambiente di elaborazione grappolo su ripristini SCSI in parecchi modi differenti. Esempi tecnici per evitare ripristini SCSI possono essere compresi meglio dalla

BUZZI, NOTARO &
ANTONIELLI D'OUX
S.R.L.

considerazione delle figure 2 e 3.

La figura 2 rappresenta un diagramma di flusso di una forma d'attuazione di un procedimento 100 per gestire risorse di memorizzazione in un ambiente di elaborazione a grappolo. Il procedimento della figura 2 può essere implementato per mezzo di un motore di gestione di risorse che esegue un controllore di memoria connesso ad una struttura SAN. In alcune forme d'attuazione, il motore di gestione di risorse può essere eseguito su di una CPU associata ad un commutatore come il commutatore 36 della figura 1. In altre forme d'attuazione, la CPU può essere associata ad un dispositivo SAN diverso dal commutatore. Ad esempio, un motore di gestione di risorse può essere eseguito su uno o più nodi di una SAN.

Durante il funzionamento di una SAN, può essere ricevuto un comando di porta di collegamento (PLOGI). Come è noto nel settore, un comando PLOGI è un comando fibre channel in cui un nodo si collega ad un dispositivo di memorizzazione connesso ad un SAN. Un nodo può eseguire un comando PLOGI dopo che la struttura ha assegnato un ID fibre channel (S_ID) al nodo. Come è anche convenzionalmente noto, l'S_ID di un nodo può essere assegnato quando un nodo esegue un comando di collegamento alla

BUZZI, NOTARO &
ANTONIELLI D'OUX
S.R.L.

struttura (FLOGI).

Alla fase 102, possono essere estratti l'S_ID ed il WWN di un nodo di grappolo. L'estrazione può verificarsi in istanti differenti. Ad esempio, l'estrazione può verificarsi quando un nodo emette un comando PLOGI. Una volta estratti, l'S_ID ed il WWN possono essere aggiornati e possono essere memorizzati in un mezzo leggibile tramite computer. In alcune forme d'attuazione, questo mezzo leggibile tramite computer può essere parte di una SAN e può essere accessibile a diversi dispositivi della SAN.

Alla fase 104, una riserva LUN può essere mantenuta per un dato nodo. In effetti, il dato nodo può avere un diritto esclusivo di utilizzare il LUN riservato. Come menzionato sopra, i nodi di grappolo spesso comunicano fra loro utilizzando un segnale di pulsazione. Alla fase 106, un dispositivo SAN può rilevare un insuccesso nel ricevere un segnale di pulsazione tempestiva. Sebbene l'insuccesso nel ricevere un segnale di pulsazione può solo indicare un insuccesso di collegamento di comunicazione fra l'emettitore ed il ricevitore della pulsazione, l'insuccesso può causare, come illustrato alla fase 108, la determinazione che un nodo di grappolo non è funzionante.

BUZZI, NOTARO &
ANTONIELLI D'OUX
S.r.l.

Nella forma d'attuazione della figura 2, la determinazione che un nodo non è funzionante, può provocare l'emissione di un ripristino SCSI da parte di un altro nodo. Come illustrato alla fase 110, un comando di ripristino SCSI può essere inviato per rilasciare una riserva LUN mantenuta per il nodo che si crede essere non funzionante (il nodo "morto"). Alla fase 112, il comando di ripristino SCSI può essere convertito in uno scollegamento di processo di un terzo. Questa conversione può, ad esempio, essere eseguita da un motore di gestione di risorse d'esecuzione.

Alla fase 114 un comando di sconnessione per il nodo "morto" può essere inviato per conto del nodo "morto" da un terzo. Ad esempio, un motore di gestione di risorse può accedere ad un mezzo leggibile tramite computer che memorizza l'S_ID ed il WWN del nodo "morto". Il motore di gestione di risorse può utilizzare l'S_ID ed il WWN del nodo "morto" per scollegare il nodo "morto". Questa sconnessione di processo di un terzo può provocare il rilascio di una riserva LUN mantenuta per il nodo scollegato.

Come illustrato alla fase 116 della figura 2, altri nodi di un grappolo possono anche scollegare o essere scollegati e può essere iniziato

un ripristino di collegamento di protocollo di inizializzazione a ciclo (LIP). Il ripristino di collegamento LIP della fase 118 può essere seguito dalla generazione della fase 120 di una notifica di cambiamento di stato. Nella forma d'attuazione della figura 2, una notifica di cambiamento di stato può provocare l'esecuzione di un collegamento di porta e la ricerca di una riserva LUN da parte di nodi di grappolo attivi, nodi che non sono morti. Il collegamento di porta dei nodi di grappolo attivi può essere visto alla fase 122, se il nodo "morto" non era morto, esso può essere in grado di ottenere nuovamente la sua riserva LUN. Se il nodo "morto" era morto, altri nodi di grappolo possono ora essere in grado di catturare la riserva LUN mantenuta dal nodo "morto". In effetti, le risorse di memoria mantenute dal nodo morto verranno rese disponibili per i nodi "vivi" - il che comporta una migliore utilizzazione delle risorse di memorizzazione - senza un ripristino SCSI.

Un'altra forma d'attuazione di un procedimento 200 per gestire risorse di memorizzazione in un ambiente di elaborazione a grappolo può essere visto nella figura 3. Il procedimento della figura 3, come il procedimento della figura 2, può essere implementato per mezzo di

BUZZI, NOTARO &
ANTONELLI D'OLIX
S.r.l.

un motore di gestione di risorse. Questo motore può essere disposto in un qualunque numero di posizioni. Ad esempio, il motore può essere disposto in corrispondenza di un commutatore, di un nodo, o di un controllo di memorizzazione connesso ad una struttura Fibre Channel.

Come illustrato alla fase 202, il procedimento 200 può coinvolgere il ricevimento di un comando di riserva LUN SCSI. Un comando di riserva SCSI tipico può essere cancellato con un ripristino SCSI. Come menzionato sopra, i ripristini SCSI possono provocare vari problemi in un ambiente di elaborazione a grappolo. Come tale, alla fase 204, il comando di riserva SCSI può essere convertito in un comando fuori riserva persistente SCSI con un'azione di servizio di RISERVA. La conversione dalla riserva SCSI alla riserva persistente SCSI può essere eseguita, ad esempio, un'esecuzione di un motore di gestione di risorse. Il comando fuori riserva persistente può mantenere una riserva LUN persistente come illustrato alla fase 206 per il nodo di mantenimento, il nodo emettendo un comando di riserva SCSI.

Alla fase 208, può essere determinato che il nodo di mantenimento non è funzionante. In risposta a questa determinazione, può essere emesso un

GIULIO NOTARO &
ANTONELLI D'OUX
s.r.l.

comando di ripristino SCSI. Il comando di ripristino SCSI della fase 210 può essere convertito alla fase 212 in un comando di riserva persistente SCSI con un'azione di servizio di CANCELLAZIONE. Nel funzionamento, il comando di riserva persistente SCSI con un'azione di servizio di CANCELLAZIONE può rilasciare la riserva LUN mantenuta dal comando fuori riserva persistente SCSI iniziale. Il rilascio LUN della fase 214 può effettivamente rilasciare risorse di memorizzazione mantenute da nodi ritenuti essere non funzionanti alla fase 208. Questo può provocare una migliore utilizzazione delle risorse di memorizzazione in un ambiente di elaborazione a grappolo, e la migliore utilizzazione può essere realizzata senza utilizzare ripristini SCSI.

Vari cambiamenti delle forme d'attuazione precedenti sono contemplati dalla presente descrizione. Ad esempio, forme d'attuazione della presente descrizione possono essere implementate in SAN aventi qualunque numero di topologie. Possono esservi, ad esempio, numerosi controllori di memoria, può esservi un motore di gestione di risorse che viene eseguito su ciascun nodo di un grappolo, o possono esservi singoli motori di gestione di risorse in esecuzione in ciascuna zona di un ambiente di elaborazione a grappolo.

BUZZI, NOTARO &
ANTONELLI D'OUX
S.R.L.

Sebbene le forme d'attuazione descritte siano state descritte in dettaglio, dovrebbe essere compreso che vari cambiamenti, sostituzioni ed alterazioni possono essere apportati alle forme d'attuazione senza uscire dal loro spirito e dalla loro portata.

BUZZI, NOTARO &
ANTONIELLI D'OUX
s.r.l.

RIVENDICAZIONI

1. Un procedimento per gestire risorse di memorizzazione in un ambiente di elaborazione a grappolo, il procedimento comprendendo:

il mantenimento di una riserva su di una risorsa di memorizzazione per un primo nodo dell'ambiente di elaborazione a grappolo;

l'esecuzione di uno scollegamento di processo di un terzo per il primo nodo; ed

il rilascio della riserva mantenuta per il primo nodo.

2. Il procedimento della rivendicazione 1, comprendente inoltre la determinazione che il primo nodo non è funzionante.

3. Il procedimento della rivendicazione 1, comprendente inoltre:

un insuccesso nel ricevere un segnale di pulsazione del primo nodo; e

la determinazione che il primo nodo non è funzionante a seguito dell'insuccesso nel ricevere il segnale di pulsazione.

4. Il procedimento della rivendicazione 1, comprendente inoltre l'iniziare lo scollegamento di processo di un terzo da un commutatore associato all'ambiente di elaborazione a grappolo.

6. Il procedimento della rivendicazione 1,

BUZZI, NOTARO &
ANTONELLI D'OUX
s. r. l.

comprendente inoltre la creazione di una zona nel sistema di memorizzazione Fibre Channel, la zona comprendendo il primo nodo dell'ambiente di elaborazione a grappolo, un secondo nodo dell'ambiente di elaborazione a grappolo, ed almeno una risorsa di memorizzazione.

7. Il procedimento della rivendicazione 1, in cui la risorsa di memorizzazione comprende un'unità di disco rigido.

8. Il procedimento della rivendicazione 1, comprendente inoltre;

informazioni di identificazione di memorizzazione riguardo al primo nodo; e

l'impiego di informazioni di identificazione per emettere una richiesta di scollegamento di processo di un terzo.

9. Il procedimento della rivendicazione 1, comprendente inoltre:

l'assegnazione di un nome world wide e di un identificatore di sorgente per il primo nodo; e

l'impiego del nome world wide e dell'identificatore di sorgente per emettere una richiesta di scollegamento di processo di un terzo per conto del primo nodo.

10. Il procedimento della rivendicazione 1, comprendente inoltre:

LUZZI, NOTIARO &
ANTONELLI D'OLIX
s.r.l.

l'assegnazione di un nome world wide e di un identificatore di sorgente per il primo nodo;

l'assegnazione di un secondo nome world wide e di un identificatore di sorgente per un secondo nodo;

l'impiego del nome world wide e dell'identificatore di sorgente per emettere una richiesta di scollegamento di processo di un terzo per il primo nodo;

lo scollegamento del primo nodo; e

lo scollegamento del secondo nodo.

11. Il procedimento della rivendicazione 1, in cui il primo nodo comprende un server.

12. Il procedimento della rivendicazione 1, in cui il primo nodo comprende una stazione di lavoro.

13. Il procedimento della rivendicazione 1, comprendente inoltre lo scollegamento di un secondo nodo dell'ambiente di elaborazione a grappolo.

14. Il procedimento della rivendicazione 13, comprendente inoltre la generazione di una notifica di cambiamento di stato con un commutatore associato all'ambiente di elaborazione a grappolo.

15. Il procedimento della rivendicazione 13, comprendente inoltre:

l'inizio di un ripristino di collegamento di protocollo d'inizializzazione a ciclo; e

BUZZI, NOTARO &
ANTONELLI D'OLIX
S.R.L.

la generazione di una notifica di cambiamento di stato in risposta al ripristino di collegamento di protocollo d'inizializzazione a ciclo.

16. Il procedimento della rivendicazione 13, comprendente inoltre:

in risposta alla notifica di cambiamento di stato, l'esecuzione di un collegamento di porta con il secondo nodo.

17. Un procedimento per gestire risorse di memorizzazione in un ambiente di elaborazione a grappolo, il procedimento comprendendo:

la ricezione di un comando di riserva di interfaccia di sistema di un piccolo computer per cercare di riservare una risorsa di memorizzazione per un nodo dell'ambiente di elaborazione a grappolo; e

in risposta al comando di riserva, l'emissione di un comando fuori riserva persistente di interfaccia di sistema di piccolo computer con un'azione di servizio di riserva per riservare la risorsa di memorizzazione per il nodo.

18. Il procedimento della rivendicazione 17, in cui un'unità a mini-porta riceve il comando di riserva ed emette il comando fuori riserva persistente.

BUZZI, NOTARO &
ANTONELLI D'OLIA
S.R.L.

19. Procedimento della rivendicazione 17, comprendente inoltre il rilascio di una riserva mantenuta per il nodo mediante l'emissione di un comando fuori riserva persistente di interfaccia di sistema di piccolo computer con un'azione di servizio di cancellazione.

20. Un sistema di computer, comprendente:

un primo nodo di un ambiente di elaborazione a grappolo;

un secondo nodo dell'ambiente di elaborazione a grappolo; e

un motore di gestione di risorse azionabile per convertire un comando di ripristino di interfaccia di sistema di un piccolo componente in un comando di rilascio di risorsa di memorizzazione.

21. Il sistema della rivendicazione 20, in cui il comando di rilascio di risorsa comprende uno scollegamento di processo di un terzo.

22. Il sistema della rivendicazione 20, in cui il comando di rilascio di risorsa comprende un comando fuori riserva persistente di interfaccia di sistema di un piccolo componente con un'azione di cancellazione.

23. Il sistema della rivendicazione 20, comprendente inoltre:

un mezzo leggibile tramite computer che

BUZZI, NOTARO &
ANTONELLI D'OLIA
s.r.l.

memorizza il motore di gestione di risorse; e

un'unità di elaborazione centrale accoppiata in modo comunicativo al mezzo leggibile tramite computer ed azionabile per eseguire il motore di gestione di risorse.

24. Il sistema della rivendicazione 23, comprendente inoltre:

una pluralità di piattaforme di elaborazione accoppiate in modo comunicativo al primo nodo;

un commutatore Fibre Channel accoppiato in modo comunicativo al primo nodo; e

una pluralità di dispositivi di memorizzazione accoppiati in modo comunicativo con il commutatore Fibre Channel.

25. Il sistema della rivendicazione 24, in cui il commutatore Fibre Channel comprende l'unità d'elaborazione centrale.

Ing. Franco BUZZI

Per iscritto ALBO 289

(in proprio o per gli altri)


CCIAA
Tomo

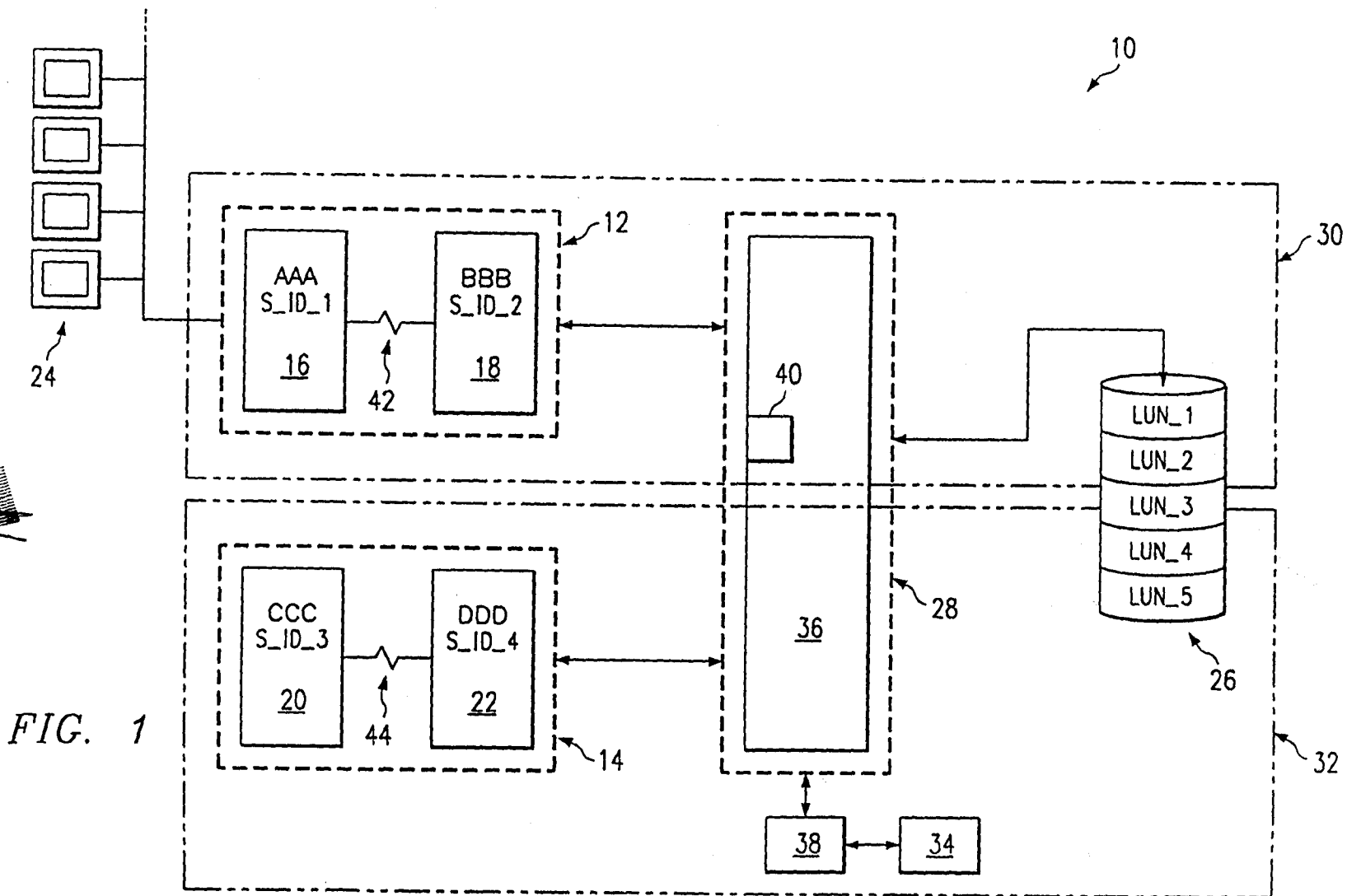
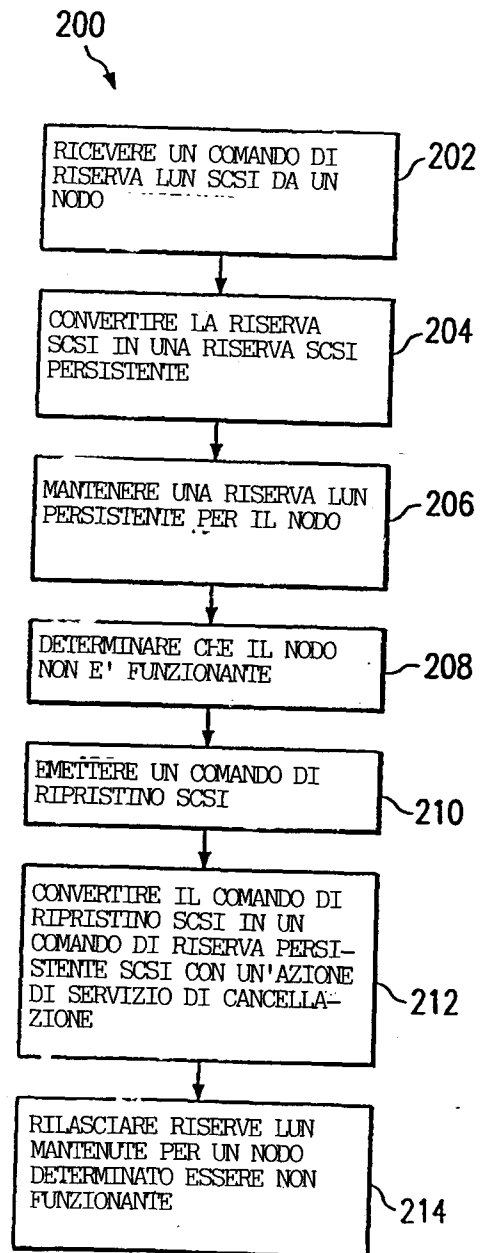
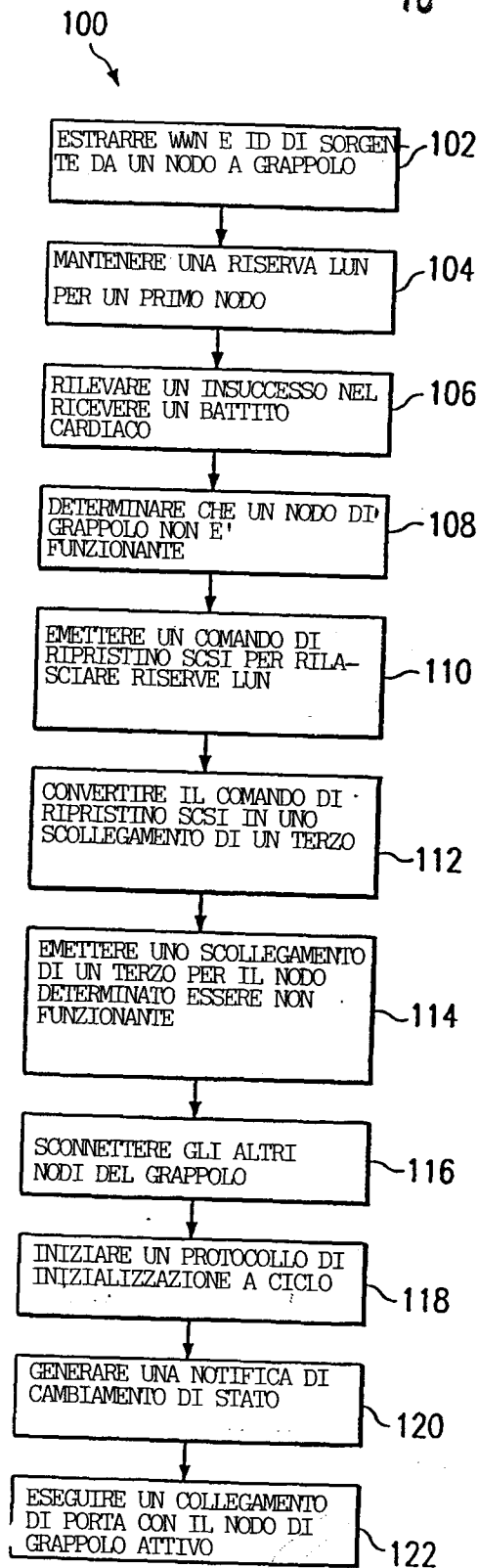


FIG. 1

C. C. A. A.
T. O. R. I. O.

Ing. Franco Rizzoli
Attestato ALBO 259
(in proprio e per gli altri)



G.C.I.A.A.
Torino